

ИНФОРМАЦИЯ И ИНФОРМАТИКА

Вводная лекция по курсу “Информатика”

1.1. Данные и информация

Мы живем в материальном мире. Все, что нас окружает, и с чем мы сталкиваемся, относится либо к физическим телам, либо к физическим полям. Физические объекты движутся и изменяются. Все это сопровождается обменом энергией и переходом ее из одной формы в другую. Все виды энергообмена сопровождаются появлением сигналов. Если эти сигналы регистрируются и сохраняются, то образуются *данные*. Данные – это зарегистрированные сигналы. Данные несут в себе информацию о событиях в материальном мире.

Данные предназначены для извлечения из них информации. Например, наблюдая излучение далеких звезд, человек получает определенный поток данных, но станут ли эти данные информацией, зависит от многих причин.

Чтобы получить информацию, необходим метод преобразования данных. Под информацией понимают продукт взаимодействия данных и адекватных им методов. Иными словами, информация – есть характеристика не сообщения, а соотношения между сообщением и его потребителем. Только соединяясь с потребителем, сообщение «выделяет» информацию.

С информацией мы постоянно сталкиваемся в реальной жизни, она, наряду с такими понятиями, как вещество и энергия, является одним из основных понятий науки. Информация – это сведения, знания и сообщения, получаемые человеком из различных источников.

Обратим внимание на следующие обстоятельства:

1. *Динамический характер информации.* Информация не является статичным объектом и существует только в момент взаимодействия данных и методов. Все прочее время она пребывает в состоянии данных. Таким образом, информация существует только в момент протекания информационного процесса.
2. *Адекватность используемых методов.* Метод преобразования данных должен соответствовать обрабатываемым данным.
3. *Диалектический характер взаимодействия данных и методов.* Обратим внимание, что данные являются объективными, поскольку это результат регистрации сигналов, вызванных изменениями в материальных телах. В то же время методы преобразования данных – субъективны. В основе естественных методов лежат биологические свойства субъектов информационного процесса. Следовательно, информация возникает и существует в момент диалектического взаимодействия объективных данных и субъективных методов.

1.2. Виды и свойства информации

Все многообразие окружающей нас информации можно сгруппировать по различным признакам:

- **по области возникновения.** Информация, отражающая процессы, явления неодушевленной природы, называется элементарной или механической, процессы животного или растительного мира – биологической, человеческого общества – социальной.
- **по способу передачи и восприятия.** Информацию, передаваемую видимыми образами и символами, называют визуальной, звуками – аудиальной, ощущениями – тактильной, запахами и вкусами – органолептической, информацию, выдаваемую и воспринимаемую средствами вычислительной техники – машинной.
- **по общественному назначению** информацию можно разбить на три вида: личная (персональная), массовая и специальная.

Обратим внимание на следующие обстоятельства:

Поскольку информация – это нечто извлекаемое из набора данных потребителем, то и качество информации оценивается с его точки зрения. Качество информации – обобщенная положительная характеристика, отражающая степень ее полезности для пользователя. *Показатель качества* – одно из положительных свойств информации (с позиции потребителя). Чтобы определить важнейшие показатели качества, необходимо оценить информацию с точки зрения ее потребителя.

Потребитель на практике сталкивается со следующими ситуациями:

- часть информации соответствует его требованиям и такую информацию называют релевантной; а часть – нет; и ее называют нерелевантной;
- если информация релевантна, но ее недостаточно для нужд потребителя, то она является неполной; если полученной информации достаточно, то такую информацию называют полной;
- полученная информация несвоевременна (например, устарела);
- часть информации, признанной потребителем релевантной, может оказаться недостоверной, т. е. содержащей скрытые ошибки;
- информация недоступна;
- информация подвержена «нежелательному» использованию и изменению со стороны других потребителей;
- информация имеет неудобные для потребителя форму или объем.

Обзор приведенных ситуаций позволяет сформулировать следующие определения свойств информации.

Релевантность – способность информации соответствовать запросам потребителя.

Полнота – свойство информации исчерпывающе (для данного потребителя) характеризовать отображаемый объект или процесс.

Своевременность или актуальность – способность информации соответствовать нуждам потребителя в нужный момент времени.

Достоверность – свойство информации не иметь скрытых ошибок.

Доступность – свойство информации, характеризующее возможность ее получения данным потребителем.

Защищенность – свойство, характеризующее невозможность несанкционированного использования или изменения.

Эргономичность – свойство, характеризующее удобство формы или объема информации с точки зрения данного потребителя.

Наконец, обобщенно характеризуя качество информации, часто используют следующее определение. Логическая, адекватно отображающая объективные закономерности природы, общества и мышления – это есть научная информация. Заметим, что последнее определение характеризует не взаимоотношение «информация – потребитель», а взаимоотношение «информация – отражаемый объект». Здесь наиболее важным свойством является свойство адекватности.

Адекватность – свойство информации однозначно соответствовать отображаемому объекту или явлению. Адекватность оказывается для потребителя внутренним свойством информации, проявляющим себя через релевантность и достоверность.

В одном терминологическом ряду с понятием информации стоит понятие “знание”. **Знание** – это переработанная информация, это информация, на основании которой путем логических рассуждений могут быть получены определенные заключения или выводы. Как философская категория знание – это проверенный практикой результат познания действительности.

1. 3. Информатика: предмет и задачи

Информатика – это область человеческой деятельности, вовлекающая в свою орбиту самые разные науки – фундаментальные и прикладные, физико-математические и технические. Своими приложениями информатика проникает во все области знаний. Можно провести аналогию с энергетикой. Энергетика – это и теплотехника, и электротехника, и физика плазмы, и ядерная энергетика, и материаловедение и это далеко не все. Информатика сродни математике в том смысле, что обе они являются слугами других наук и сами не создают материальных ценностей.

Многоликость информатики позволяют выделить в ней отдельные научные формирования, важнейшими из которых являются:

- вычислительная информация – направление, включающее отображение алгоритмов на архитектуру вычислительных систем, прикладное ПО вычислительных задач и методологию численного моделирования процессов;
- машинная информатика – фундаментальные исследования принципов построения базовых элементов, архитектуры процессорных устройств, вычислительных комплексов и сетей;
- системное программирование – совокупность теоретических дисциплин, методологий и технологий по разработке ОСНОВНОЙ, трансляторов и других программных компонентов;

- искусственный интеллект – обеспечение “очеловеченного” общения с ЭВМ, распознавание образов и различные трудноформализуемые и недетерминированные задачи;
- информационные и управляющие системы – технологии автоматизированного сбора данных, их обработка и использование в управлении производством и технологическими процессами.

Информатика – это совокупность фундаментальных и прикладных научных направлений, изучающих технические, программные и алгоритмические аспекты процессов накопления, передачи и обработки данных, а также их использование в различных областях человеческой деятельности.

Коротко это можно сформулировать следующим образом: информатика – это совокупность научных направлений, изучающих процессы обработки данных с помощью средств вычислительной техники.

Можно сказать, что информатика – это все, что связано с применением ЭВМ, их разработкой и созданием программ для них.

В информатике выделяют три основных направления: техническое – *hardware*, практическое – *software*, и теоретическое – *brainware*. Они изучают:

- проектирование, разработку и использование технических средств обработки информации (*hardware*);
- программирование и использование прикладных программ (*software*);
- теорию информации (*brainware*);
- теорию алгоритмов и вычислительных методов (*brainware*);
- математическую логику и комбинаторный анализ (*brainware*).

Особое внимание в информатике уделяется вопросам взаимодействия. Для этого даже есть специальное понятие – интерфейс. Существуют аппаратный интерфейс, программный интерфейс, аппаратно-программный интерфейс. Методы и средства взаимодействия человека с аппаратными и программными средствами называют пользовательским интерфейсом.

Среди направлений информатики для практических приложений можно выделить следующие:

- архитектура вычислительных систем;
- интерфейсы вычислительных систем;
- программирование;
- преобразование данных;
- защита информации;
- автоматизация;
- стандартизация.

В информатике все жестко ориентировано на эффективность. Вопрос, как сделать ту или иную операцию, для информатики является важным, но не основным. Основным же является вопрос, как сделать данную операцию эффективно.

1.4. Краткая история информатики

Информатика как наука стала развиваться с середины прошлого столетия, что связано с появлением ЭВМ – чрезвычайно мощного инструмента для реализации процессов получения, хранения, обработки и передачи информации. И само название этой науки появилось именно тогда, как объединение двух французских слов: l'Information + l'Automatique. Историю информатики принято разбивать на два этапа: предысторию и историю.

В предыстории выделяют четыре основных этапа. Каждый из них характеризуется по сравнению с предыдущим резким возрастанием возможностей хранения, передачей и обработки информации.

Начальный этап предыстории – освоение человеком развитой устной речи. Членораздельная речь, язык стал специфическим социальным средством хранения и передачи информации.

Второй этап – появление письменности, т.е. знаковой системы (пиктограммы, иероглифы, клинопись, линейное письмо). Теперь информация накапливается, обрабатывается в форме записей на определенных материальных носителях: глиняных табличках, папирусе и т.п. Тем самым, как любой материальный объект, характеризуется формой и содержанием. Форма – это способ хранения, т.е. используемая система знаков и синтаксис. Содержание, или **семантический аспект** хранимой информации, – это смысл записи, т.е. ссылкой на что, на какой материальный объект, процесс, служит запись в той или иной форме. Что касается возможности использовать хранимую информацию, **прагматический аспект**, то она тесно связана с такой характеристикой информации как **полезность**.

С возникновением письменности резко возросли возможности по хранению информации. Человек получил искусственную внешнюю память. Организация почтовых служб позволила использовать письменность и как средство для передачи информации. Кроме того, возникновение письменности было необходимым условием для начала развития наук. С этим же этапом, по всей видимости, связано и возникновение понятия *натуральное число*.

Третий этап – книгопечатание. Книгопечатание можно смело назвать первой информационной технологией. Воспроизведение информации было поставлено на поток, на промышленную основу. Это увеличило возможности по хранению информации и повысил ее доступность.

Последний этап предыстории связан с успехами точных наук и начинающейся в то время научно-технической революцией. Этот этап характеризуется возникновением таких мощных средств связи, как радио, телефон и телеграф, к которым по завершению этапа добавилось и телевидение. Кроме средств связи появились новые возможности по получению и хранению информации – фотография и кино. К ним также можно добавить разработку методов записи информации на магнитные носители.

С разработкой первых ЭВМ принято связывать возникновение информатики как науки, начало ее истории. Для такой «привязки» имеется несколько причин. Во-первых, сам термин «информатика» появился на свет благодаря развитию вычислительной техники, и поначалу под ним понималась наука о вычислениях (первые ЭВМ использовались для проведения числовых расчетов). Во-вторых, выделению информатики в отдельную науку способствовало такое важное свойство вычислительной техники, как ***единая форма представления обрабатываемой и хранимой информации***. Вся информация, вне зависимости от ее вида, хранится и обрабатывается в двоичной форме. Компьютер в одной системе объединил хранение и обработку числовой, текстовой и аудиовизуальной информации. В этом состояла иницирующая роль вычислительной техники при возникновении и оформлении новой науки.

Второй этап истории информатики обусловлен появлением персональных компьютеров, а третий – с разработкой глобальной сети Интернет.

На сегодняшний день информатика представляет собой комплексную научно-техническую дисциплину. Информатика под своим названием объединяет довольно обширный комплекс наук, каждая из которых занимается изучением одного из аспектов понятия *информация*. Процесс сближения научных дисциплин в единую и всеохватывающую науку *информатика* идет довольно медленно и продолжается до сих пор.

1.5. Измерение информации

Какое количество информации содержится, к примеру, в тексте романа «Война и мир», во фресках Рафаэля или в генетическом коде человека? Ответа на эти вопросы наука не дает и, по всей вероятности, даст не скоро. А возможно ли объективно измерить количество информации? Важнейшим результатом теории информации является следующий вывод:

В определенных, весьма широких условиях можно пренебречь качественными особенностями информации, выразить ее количество числом, а также сравнить количество информации, содержащейся в различных группах данных.

В настоящее время получили распространение подходы к определению понятия количества информации, основанные на том, что информацию, содержащуюся в сообщении, можно трактовать в смысле ее новизны, или иначе, уменьшения неопределенности наших знаний об объекте. Одной из общепринятых точек зрения является трактовка информации как отрицательной энтропии (***негэнтропии***) материальной системы. Если под энтропией понимать меру хаоса или неопределенности, то под информацией следует понимать меру упорядоченности системы или другими словами, как меру неоднородности распределения вещества (энергии) материальной системы в пространстве и во времени. Энтропия устраняется информацией.

В окружающей действительности часто встречаются ситуации, когда существует некоторое количество событий, каждое из которых может

произойти с определенной вероятностью. Рассмотрим это на примере “бросание монеты”. Если бросить монету на ровную поверхность, то реализуется одно из двух состояний – либо монета упадет в положение “орел” или в положение “решка”.

Перед броском существует неопределенность, как упадет монета предсказать невозможно. После броска реализуется полная определенность, так как монета займет определенное положение. При бросании четырехгранной пирамиды существует четыре возможных исхода, а при бросании кубика – шесть. Совершенно очевидно, что чем большее количество возможных событий, тем больше начальная неопределенность результата и, главное, тем большее количество информации будет получено после проведения опыта.

За единицу количества информации принято такое количество, которое мы получаем при уменьшении неопределенности в два раза. Такая единица в теории информации называется **бит**. Другими словами, **бит** в теории информации – это количество информации, необходимой для различения двух равновероятных сообщений. Если вернуться к опыту с бросанием монеты, то здесь неопределенность как раз уменьшается в два раза (из двух возможных событий реализуется одно) и, следовательно, полученное количество информации после выполнения броска равно одному биту.

В компьютере информация представлена в двоичном коде, алфавит которого состоит из двух цифр (0 и 1). Эти цифры можно рассматривать как два равновероятных состояния. При записи одного двоичного разряда реализуется выбор одного из двух возможных состояний (одной из двух цифр) и, следовательно, один двоичный разряд несет количество информации в один бит. Соответственно два двоичных разряда несут информацию два бита, три разряда – три бита и т.д. Количество информации в битах равно количеству двоичных разрядов. Обратим задачу и поставим такой вопрос: “Какое количество различных двоичных чисел N можно записать с помощью I двоичных разрядов?” С помощью одного двоичного разряда можно записать два различных числа: 0 и 1, т.е. $N=2=2^1$. С помощью двух двоичных разрядов можно записать четыре двоичных числа: 00, 01, 10, и 11, т.е. $N=4=2^2$. С помощью трех двоичных разрядов – восемь двоичных чисел, т.е. $N=8=2^3$. В общем случае количество различных двоичных чисел можно определить по формуле $N=2^I$.

Мы получили очень важную формулу, которая связывает между собой количество возможных исходов N и количество информации I . Если известно количество возможных исходов и необходимо определить I , то формула принимает вид показательного уравнения относительно I , которое дает решение

$$I = \log_2 N.$$

Впервые эту формулу получил американский инженер Р. Хартли в 1928 г. Измерение информации с помощью полученной формулы в теории информации носит название энтропийного подхода. Этот способ измерения исходит из следующей модели. Получатель информации имеет определенное

представление о возможных наступлениях некоторых событий. Эти представления в общем случае недостоверны и выражаются вероятностями, с которыми он ожидает то или иное событие. Общая мера неопределенности (энтропия) характеризуется некоторой математической зависимостью от совокупности этих вероятностей. Количество информации в сообщении определяется тем, насколько уменьшится эта мера после получения сообщения.

Поясним эту идею на простом примере. Пусть имеется колода из 32 карт. Возможностей выбора одной из карт из колоды столько, сколько карт в колоде – 32. Шансы выбрать некоторую определенную карту одинаковы ($p=1/32$). Произведя выбор, мы устраняем неопределенность. Если теперь определить количество информации как меру устраненной неопределенности, то по формуле Хартли получаем

$$I = \log_2 32 = 5.$$

Полученное значение имеет интересную интерпретацию. Она характеризует число двоичных вопросов, ответы на которые позволяют выбрать либо “да”, либо “нет”. Например, если мы вытянули даму пик, то такими вопросами будут:

- | | | |
|-----------------------------|-----|---|
| 1. Карта красной масти? | Нет | 0 |
| 2. Трефы? | Нет | 0 |
| 3. Одна из четырех старших? | Да | 1 |
| 4. Одна из двух старших? | Нет | 0 |
| 5. Дама? | Да | 1 |

Этот выбор можно описать последовательностью из пяти двоичных символов 00101.

Определим теперь, являются ли равновероятными сообщения “первой выйдет из дверей здания женщина” и “первым выйдет из дверей здания мужчина”. Однозначно ответить на этот вопрос нельзя. Все зависит от того, о каком именно здании идет речь. Если это, например, станция метро, то вероятность выйти из дверей первым одинакова для мужчины и женщины, а если это военная казарма, то для мужчины эта вероятность значительно выше, чем для женщины. Для задач такого рода американский ученый Клод Шеннон предложил в 1948 г. другую формулу определения количества информации, учитывающую неодинаковую вероятность сообщений в наборе:

$$I = - \sum_{i=1}^N p_i \log_2 p_i,$$

где I – количество информации, N – количество возможных событий, p_i – вероятности отдельных событий.

Для частного, но широко распространенного случая, когда события равновероятны (т.е. $p_i = 1/N$), величина количества информации I принимает максимальное значение

$$I = - \sum_{i=1}^N \frac{1}{N} \log_2 \frac{1}{N} = \log_2 N,$$

и формула Шеннона превращается в формулу Хартли.

1.6. Информационные процессы

Все действия, связанные с получением, передачей, хранением и переработкой информации, называются информационными процессами

Человек постоянно имеет дело с информационными процессами в ходе своей эволюции. Любая деятельность, любое взаимодействие индивида с окружающим миром требует информационного обеспечения (информационных процессов). Информация используется для устранения неопределенности принятия решения относительно любого действия, поступка.

Процесс передачи информации от человека к человеку двусторонний: есть источник и есть приемник информации. В настоящее время информация может быть передана с помощью технических средств связи, таких как телефон, радио, телевидение, Интернет. Эти технические средства называются каналами передачи информации. Количество информации, передаваемой за единицу времени по информационному каналу, называется скоростью передачи информации или скоростью информационного потока.

Процесс хранения информации заключается в том, что человек запоминает информацию в собственной памяти, либо заносит ее на внешний носитель информации.

Процесс обработки информации называется ее преобразование по каким-либо правилам или законам.

Основной и постоянной функцией мозга и нервной системы человека или животного является преобразование информации о состоянии окружающей среды и выбор наиболее целесообразного поведения. Процесс преобразования исходной информации в информацию, отражающую результат решения какой-либо задачи – это и есть решение задачи, поставленной перед человеком в любом виде его деятельности.

В качестве примера рассмотрим хрестоматийную задачу – решение квадратного уравнения. Такого вида уравнения отражают фундаментальные законы в целом ряде предметных областей. Прежде всего необходимо получить исходные данные, значения коэффициентов A , B , C . Отсутствие хотя бы одного из них приведет к тому, что решение получить нельзя – исходная информация неполная. Располагая полной исходной информацией, порождаем информационный процесс – получение новой, искомой информации. На первом этапе требуется определить множество, которому принадлежит искомое решение: поле комплексных чисел или поле вещественных чисел. Чтобы устранить возникшую неопределенность (**энтропию процесса**) относительно выбора множества возможных решений требуется дополнительная информация! **Энтропия** устраняется **информацией!** Эту дополнительную информацию мы можем получить либо извне, либо самостоятельно на основе исходных данных (тех, которыми уже располагаем). Из школьного курса математики хорошо известно, что дополнительная информация, требуемая для устранения неопределенности с выбором множества возможных решений, легко может быть получена на основе исходных данных (путем вычисления дискриминанта). На следующем

этапе, используя информацию в виде формул для корней квадратного уравнения, получаем решение и затем организуем вывод пользователю результатов процесса.

1.7. Алгоритм и его свойства

Алгоритм – одно из основных понятий информатики, которое не имеет строгого определения. Близкими по смыслу словами являются слова “рецепт”, “метод”, “программа”. С понятием “алгоритм” тесно связано понятие “исполнитель” алгоритма. В качестве такового может выступать человек, робот, компьютер, станок и т.д. Исполнитель “знает” и “умеет” выполнять набор команд, называемый системой команд исполнителя – СКИ. Подчеркнем, что исполнитель действует формально, в рамках своей системы команд.

Как определение, можно дать следующее: Алгоритм – понятная и точная совокупность действий, предписанная исполнителю и направленная на достижение поставленной цели. Другими словами, алгоритм – это набор инструкций, который описывает, как некоторое задание может быть выполнено.

Разберем простой алгоритм, который назовем алгоритмом умножения. Он описывает процесс умножения положительных целых величин A и N , в результате чего получается число M . Целью его, как и всех алгоритмов, является описание некоторого задания в терминах других заданий, которые предварительно определены, либо достаточно просты, чтобы быть понятными без определений.

Шаг 1. Присвоить I значение 0 (задать переменной I значение равное 0).

Шаг 2. Присвоить P значение 0.

Шаг 3. Если N равно 0 или A равно 0, то перейти к шагу 7.

Шаг 4. Присвоить P значение $P+A$.

Шаг 5. Присвоить I значение $N+I$.

Шаг 6. Если I меньше N , то перейти к шагу 4.

Шаг 7. Присвоить M значение P .

Шаг 8. Конец.

Рассмотрим пять основных свойств или атрибутов алгоритма: наличие входа и выхода (результативность), определенность, выполнимость и конечность. Так, чтобы алгоритм умножения начал работу, нужно задать числовые значения A и N – это и есть вход алгоритма. Вычисляемая величина M является его выходом или результатом. В хорошо разработанном алгоритме имеются два четко различимых самостоятельных множества данных, образующие вход и выход алгоритма. Элементы этих двух множеств называются параметрами или аргументами. Лучше использовать термин “параметр”, поскольку он применяется для описания объекта, который остается неизменным при данном обращении к алгоритму, но который может изменяться при других обращениях.

Начальным этапом разработки алгоритма является определение элементов множества входных и выходных параметров. На втором этапе формируется интуитивная идея связи между этими двумя множествами. На заключительном этапе интуитивная идея формулируется в виде формального набора инструкций, которые описывают указанную связь. Этот формальный набор инструкций и является алгоритмом.

Для того чтобы алгоритмом можно было пользоваться, каждая из инструкций, о которых говорилось выше, должна удовлетворять двум условиям: определенности и выполнимости.

Под **определенностью** понимается, что инструкция определена четко и недвусмысленно: не должно быть никаких сомнений относительно того, выполнение каких операций включает в себя инструкция.

Пример: *или...*

Выполнимость: говорят, что инструкция выполнима, если включенные в нее операции достаточно элементарны, чтобы их за конечное время мог выполнить человек, вооруженный карандашом и бумагой.

Пример: присвоить X значение, равное наибольшему вещественному числу, меньшему 1.

Обеспечить выполнимость – значит исключить из алгоритма невыполнимые инструкции; обеспечить определенность – значит исключить неясные или бессмысленные инструкции.

В отличие от двух последних свойств **конечность** является свойством алгоритма в целом, а не отдельных его инструкций. Приведем пример алгоритма, который никогда не заканчивается:

Шаг 1. Присвоить I значение 1.

Шаг 2. Присвоить I значение $I+1$.

Шаг 3. Перейти к шагу 2.

Конечность алгоритма умножения доказывается его прокруткой, т.е. выполнением всех его инструкций на бумаге.

К дополнительным свойствам алгоритма относят ***дискретность, массовость и эффективность***.